

# Relative Utility of Computerized Versus Paper-and-Pencil Tests for Predicting Job Performance

Charles H. Cory

Navy Personnel Research and Development Center

This article, the second of two, presents predictive validity data for on-job performance for a set of computerized Graphic and Interactive Processing (*GRIP*) tests in conjunction with data for both experimental paper-and-pencil and operational tests. Validity coefficients for job element and global criteria are reported for four different jobs. Experimental variables substantially enhanced the predictive accuracy of the operational battery for Sonar Technicians. Most experimental tests with significant validities were computer-administered. The *GRIP* tests were more useful than paper-and-pencil tests for identifying personnel skilled in *Interpreting Visual Displays*, *Adjusting Equipment*, and *Working Under Distractions*. They were useful supplements to paper-and-pencil tests for identifying skill in four additional job elements.

Computerized testing, in addition to its adaptive potential (Lord, 1970; Weiss, 1974; Weiss & Betz, 1973; Wood, 1974), provides stimulus characteristics and formats which frequently are closer to those found in jobs than those provided by paper-and-pencil tests. A battery of tests designed to use the capabilities of computer terminals for pictorial display and movement and interactive presentation of data, was described in a previous article (Cory, Rimland, and Bryson, 1977) and labeled the Graphic and Interactive Processing (*GRIP*) series.

The *GRIP* battery was composed of eight computer-administered tests, each designed to measure one of five abilities as defined by Mecham and McCormick (1969). The five abilities—Short Term Memory, Perceptual Speed, Closure, Movement Detection, and Dealing with Concepts/Information—are shown below in conjunction with the tests which were designed to measure them.

Test	Attribute
1. <i>Memory for Objects, Memory for Words and Memory for Numbers (Visual)</i>	Short Term Memory
2. <i>Comparing Figures</i>	Perceptual Speed
3. <i>Recognizing Objects</i>	Closure
4. <i>Memory for Patterns</i>	Movement Detection
5. <i>Twelve Questions and Password</i>	Dealing with Concepts/Information

In addition, three latency measures and a measure of total accuracy of pattern matching were derived using the special measurement capabilities of the computerized equipment. WORD-LAT and FIG-LAT consisted of total elapsed time before initiation of the testee's responses to *Memory for Words* and *Comparing Figures*, respectively. CLO-LAT measured the total time required to identify the mutilated pictures presented in the *Recognizing Objects* test,

---

APPLIED PSYCHOLOGICAL MEASUREMENT  
Vol. 1, No. 4 Fall 1977 pp. 551-564  
© Copyright 1977 West Publishing Co.

and PAT-ERR measured the total distance by which the responses missed the location of the stimuli in *Memory for Patterns*.

Two sets of comparison tests were used: (1) a battery of alternate tests (the *A* battery), largely paper-and-pencil, selected to provide close approximations to the computerized measures, and (2) the battery of tests used operationally (the *O* battery) for personnel selection and classification in the Navy.

The *A* battery included the following six tests from the *ETS Kit of Reference Tests for Cognitive Factors* (French, Ekstrom, & Price, 1963): *Object Number*, *Gestalt Completion*, *Concealed Words*, *Hidden Patterns*, *Nonsense Syllogisms*, and *Inference*. Also the *A* battery contained *Memory for Numbers* (Auditory), a digit-span test; *Drift Direction*, a motion-picture test measuring ability to detect minimal movements of a cursor; and *Counting Numbers*, a paper-and-pencil test measuring perceptual speed.

The *O* battery included the *Armed Forces Qualification Test* (AFQT), the *General Classification Test* (GCT), the *Arithmetic Reasoning Test* (ARI), the *Mechanical Test* (MECH), the *Clerical Test* (CLER), the *Sonar Pitch Memory Test* (SONR), the *Radio Code Aptitude Test* (RADO), the *Electronic Technician Selection Test* (ETST), and *Shop Practices* (SHOP). The AFQT measures verbal, quantitative and spatial visualization abilities and tool knowledge, and GCT measures vocabulary and verbal reasoning. The titles of the other tests indicate their content areas.

A principal components analysis of the inter-correlation matrix for the *A* and *O* batteries and for the computerized variables (except for the four special scores) identified seven factors: General Intelligence, Perceptual Speed, two Short Term Memory factors, Technical Knowledge, Closure, and Sequential Reasoning. Sequential Reasoning—a factor found only in the computerized battery—was defined essentially by the interactive data processing characteristics of *Twelve Questions* and *Password*. The computerized and paper-and-pencil measures of Closure—*Recognizing Objects* and *Gestalt Comple-*

*tion*—differed substantially in factor structure and were not substantially correlated.

Thus, the principal components analysis did not indicate that a movement detection attribute was being measured. It indicated, however, that there were differences between the paper-and-pencil and computerized measures of Closure. Both computerized and paper-and-pencil tests measured Short Term Memory, but the computerized measures supplied the bulk of the measurement of the attribute. There were no substantial differences in the loadings of the computerized and the paper-and-pencil measures of Perceptual Speed.

## Method

### Data Collection

The *A* and the *GRIP* batteries were administered to students at the Naval Training Center in San Diego during May and June of 1972. Students were chosen from personnel in the first two weeks of technical training for Electrician's Mate (EM), Personnelman (PN) and Sonar Technician (ST)—three jobs having widely varied duties. Personnel in their final week of recruit training who had not yet been designated for *A*-school training were also tested. Most of the personnel in this latter group were subsequently assigned to apprenticeship ratings—relatively unskilled entrance-level types of positions consisting largely of routine custodial and maintenance tasks.

Ten to eleven months subsequent to the testing, at which time students should have finished training and been assigned to jobs for several months, supervisory ratings of on-job performance were collected by mailed questionnaire. These evaluations were done by the immediate supervisors of the research subjects, consisting for the most part of Petty Officer personnel. Sample sizes for the test administration and the on-job follow-up are shown in Table 1.

Returns were received for 65% of the mailed sample. 24% of the questionnaires could not be completed because of transfer, school assign-

Table 1  
Sample sizes for Test Administration and Performance Followup

Occupational Group	Test Administration	Performance Followup	Unavailable because of Transfer, School Assignment Discharge, etc.	Not Returned
Personnelman (PN)	61	54	4	3
Electrician's Mate (EM)	77	27	33	17
Sonar Technician (ST)	45	37	5	3
Recruits in the upper half of the recruit distribution of ability (Classified as Undifferentiated Apprenticeships)	172 355 <sup>a</sup> 100%	111 229 65%	44 86 24%	17 40 11%

<sup>a</sup>30 cases were dropped because of incomplete experimental data, bad records or rating misclassifications.

ment, discharge or similar reason. 11% of the sample was neither present nor accounted for. Thus, despite a tracer follow-up which was mailed subsequent to the original follow-up, criterion information was not collected for about 35% of the original sample. Transfers accounted for about 5% of this figure. Because of the study's tight time schedule, no attempt was made to locate personnel who had been transferred from the command shown on the source tape.

The majority of the 33 Electrician's Mate personnel who are shown as unavailable in Table 1 had been accepted for a nuclear training course lasting almost a year, and consequently had not receive job assignments by the time of the follow-up. Of the 44 Apprenticeship personnel shown as unavailable in Table 1, more than one-fourth had been separated and another one-fourth were attending special training courses at school.

### Criterion Ratings

Ratings of on-job performance were collected using an adaptation of the *Position Analysis Questionnaire* (PAQ), a broad-based empirically derived instrument which has been extensively used for job classification research (see McCormick, Jeanneret, & Mecham, 1972). The adapted questionnaire was used to rate overall job performance as well as performance on the 42 elements which were judged by a panel of Chief Petty Officers as relevant to one or more of the jobs. All ratings were on the 5-point scales used for the original PAQ, with scores 1 to 5 representing the lowest 5%, next lowest 25%, middle 40%, next highest 25%, and highest 5%, respectively. The types of elements included in positions varied across jobs.

### Analysis

Validity coefficients were computed within jobs for any element having an  $N$  of 20 or more. Then, 12 job elements having substantial numbers of significant validities were selected for

further analysis. Subsample sizes for these elements are shown in Table 2.

For each job, the validity coefficients for each element which were statistically significant from zero were identified. Then, the significant validity coefficients in the *A*, *O*, and *GRIP* batteries were compared in terms of overall numbers and the types of personal attributes which were measured by the tests. Similar analyses were carried out for the overall performance criterion.

Multiple regression analyses were computed to identify the subset of the *O* battery that was most predictive of the shrunken multiple  $R$  for global performance. In addition, increments that *GRIP* and *A* variables supplied to the shrunken multiple regression validity coefficient of the *O* variables were computed. The multiple regression analyses were carried out by means of a step-wise predictor selection technique, with an accretion algorithm. Selection of variables was stopped at the point at which the shrunken multiple validity coefficient dropped. The shrunken values were computed by  $F$  ratios using a technique recommended by Darlington (1968, Formula 14, p. 174).

## Results

### Validity for Job Elements

The significant zero-order validity coefficients of operational variables for the 12 job elements are shown in Table 3. Those for the *A* and the *GRIP* tests are shown in Table 4. In these tables, rows with no statistically significant  $r$ 's have been omitted. Cells not containing job elements are blank, and cells having nonsignificant validities are marked with dashes.

Tables 3 and 4 reveal the tests which were predictive of each job element. In these tables, 13 of the significant correlations of the test scores were negative, while two significant correlations of the special scores were positive. Since, in the absence of explanatory hypotheses, these correlations were opposite to the direction which would be expected, they were interpreted conservatively as being significant as a result of

Table 2  
Sample Sizes for Twelve Job Elements

Job Element	EM (27)	PN (54)	ST (37)	UA (111)
Using Written Materials		48	30	71
Using Pictorial Materials	20		32	
Using Visual Displays			35	66
Using Spoken Verbal Communication	20	52	36	92
Manual Control-Non-precision Tools	27			80
Adjusting Machines/Equipment	23		29	
Persuading/Influencing Others		40		
Exchanging Routine Information		51		69
Unusually Good Precision			29	69
Attention to Details, Completing Work	25	51	36	102
Working with Distractions		48		84
Keeping up to Date		52	30	86

Table 3  
Significant Zero-Order Validities of the Operational Variables  
For Performance Marks for Twelve Common Job Elements

Predictor Variable	Occupat. Group	Skill Writ	Pict Mats	Job Element								Work Dist	Up to date
				Vis Displ	Verb Com	N-Prec. Tools	Adj Equip	Influ Othrs	Rout Info	Good Prec	Attn Dtails		
AFQT	ST	--	--	-33*	--		--				--	--	--
GCT	PN	--			50**			--	--	--	--	--	36*
	UA	--			22*	--				27*	--	--	--
ARI	ST	49**	--	--	--		--				--	--	--
	UA	--		--	24*	--				41*	27*	23*	--
MECH	PN	--			--			55**	--			38*	--
CLER	UA	25*		--	20*	--				--	29*	30**	--
SONR	PN	--		--	--			--	--	--	--	--	37*
	UA	-26*		--	--	--			--	--	--	--	-26*
RADO	EM		--*	--	--	--	--					-44*	
	ST	36*		--	33*	--	--				--	37*	39*
	UA	--		--	22*	--				--	--	--	--
ETST	UA	--		--	--	--				28*	--	--	24*
SHOP	PN	--			45*			42*	--			--	--

Column N's

EM		15		16	21	18				19			
PN	26		29	27			20	28		29	27		31
ST	29	30	33	34		27			27	34			28
UA	69		66	90	79			67	67	100			82

- Note:
1. Decimal points omitted from validity coefficients.
  2. A blank cell indicates missing data.
  3. A -- (double dash) indicates nonsignificant validity.

\*  $p < .05$   
\*\*  $p < .01$

Table 4  
Significant Zero-Order Validities of the Experimental Variables  
for Performance Marks for Twelve Common Job Elements

Job Element													
Predictor Variable	Occupat. Grp.	Skill Writ	Pict Mats	Vis Displ	Verb Com	N-Prec Tools	Adj Equip	Influ Othrs	Rout Info	Good Prec	Attn Dtails	Work Dist	Up to Date
<u>A Battery</u>													
Obj. No.	EM		--		--	--	--				-38*		
Mem. for Nos. (A)	EM PN	29*	--		-- 42**	--	--	--	--		-42*	--	--
Counting Numbers	UA ST	--	--	--	--	--	--		--	--	26**	--	39*
Gest Comp	EM		--		--	-40*	--				-42*		
Hidden Patterns	EM PN ST	--	-- 39*	--	-- 30**	-36*	--	32**	--	--	--	--	34* 35*
N-sns Syls.	EM		--		--	--	--				-38*		
Inf.	PN	37*			42**			49**	37*		--	--	29*
<u>GRIP Battery</u>													
Mem. Obj.	FM UA	--	--	--	--	-24*	-53**		--	--	--	-24*	--
Mem. Words	ST PN	42* 29*	--	--	--		--	--	--	--	--	--	--
Mem. for Nos. (V)	ST PN	40* 33*	--	--	-- 48**		--	--	--	--	--	--	-- 33*
Comp. Figs., m.p.	UA PN	25* --	--	--	20* --	--	--	--	--	29*	30**	--	26* 44**
Rec. Objs	ST	45*	--	--	--		--			--	--	--	--
Mem. for Pats. f.r.	ST UA PN	40* -- --	--	--	--	--	--	--	--	--	24*	--	38* 28** 25*
12 Ques.	PN ST	42** --	--	--	44** --		--	36*	55**	--	--	42**	31* 37*
Password	PN ST	33* --	--	--	43** --		46*	--	--	--	30*	--	--
WORD-LAT	EM		--		--	--	-41*				--		
CLO-LAT	ST	--	-47**	-44**	--		--			--	--		-37*
PAT-ERR	UA PN	-- --		--	--	--	--	--	--	--	-21*	--	-23* --
FIG-LAT	PN ST UA	-- -- --	--	34*	-- --	--	--	-36*	--	--	--	--	--
Column N's													
EM	PN	45	20		20	27	23	37	47		25	44	48
ST	PN	29	31	34	36		29			29	47	84	30
UA	PN	71		66	92	80			69	69	102		86

Note:  
1. Decimal points omitted from validity coefficients.  
2. A blank cell indicates missing data.  
3. A -- (double dash) indicates nonsignificant validity.  
\* p < .05  
\*\* p < .01

chance variation. Therefore, they were not included in the discussion and interpretation of the findings. Similarly, the two negative correlations in Table 6 were excluded from the discussion and from the predictive batteries used for Table 7.

The paper-and-pencil tests of the *O* battery were primarily predictive of performance in the Verbal Communication tasks and to a lesser extent of Skill in Writing, Attention to Details, and Keeping Up-to-Date. Tests in both the *GRIP* and the *A* batteries also predicted these elements. In addition, tests in the *GRIP* battery were predictive of Visual Displays, Adjusting Equipment and Working with Distractions, elements for which none of the operational tests were predictive.

Scores for computerized and paper-and-pencil tests, designed to measure the same ability, had previously been found to have high intercorrelations (Cory, et al., 1977). Similarly, in Tables 3 and 4, the computerized and paper-and-pencil tests for the same attribute frequently have statistically significant validity coefficients for the same job element. Thus, Closure tests in both the *A* and the *GRIP* batteries were predictive of Pictorial Materials, and the Perceptual Speed tests of the *A* and the *GRIP* batteries (Counting Numbers and Comparing Figures) tended to be predictive for the same elements. The job elements for which Auditory Memory for Numbers (in the *A* battery) was significantly valid were about the same as those for which the corresponding *GRIP* test, Visual Memory for Numbers, was significantly valid. The correspondence between the significant validity coefficients of tests in the *A* and the *GRIP* batteries which measure a Dealing with Concepts/Information ability (12 Questions and Inference) was particularly striking.

The coefficients in Tables 3 and 4 are summarized in Table 5, which shows the number of significant validity coefficients in the *O*, *A*, and *GRIP* batteries for the elements in each job. More of the computerized tests than of the *O* and *A* tests were useful for identifying good performers in the following job elements: Skill in

Writing, Interpreting Visual Displays, Adjusting Equipment, Attention to Details, and Working with Distractions. Both the *GRIP* and the *A* batteries had as many significant predictors of Pictorial Materials, and Influencing Others, as did the *O* battery.

Generally, operational variables were less effective for predicting performance on elements in specific technical jobs than in apprenticeship jobs. For the three technical occupations combined, 13 correlations for *O* variables were statistically significant, compared with 14 significant correlations for the Apprenticeship group alone. In contrast, tests in the *A* and the *GRIP* batteries were more predictive of elements in the Personnelman and Sonar Technician jobs than of those in apprenticeship jobs. None of the batteries was very predictive for Electrician's Mates.

### Validity for Overall Performance

*Zero-order.* The zero-order validity coefficients of the *A*, *GRIP*, and *O* variables for global performance are shown in Table 6. Seven of the 92 coefficients for experimental variables were statistically significant. Of these, five were significant for computerized tests. In comparison, 5 of the 36 coefficients of the operational tests were statistically significant.

A total of six experimental tests, compared with two operational tests, significantly predicted overall performance in technical occupations. For the Electrician's Mate, Personnelman, and Sonar Technician jobs, respectively, 0, 1, and 5 coefficients for experimental tests had statistically significant validities compared with 0, 1, and 1 coefficients from the *O* battery which were statistically significant for these jobs.

Thus, the general predictiveness of global marks was similar to that of job elements. None of the batteries had tests which were significantly predictive for Electrician's Mates (EMs); both the *GRIP* and the *O* batteries were predictive for Personnelman (PNs). The *GRIP* battery was the most predictive for Sonar Technicians

Table 5  
 Number of Significant Zero-Order Validities of Operational and  
 Experimental Variables for Twelve Job Elements

Job Elements	EM		PN		ST		UA		
	0	A	0	A	0	A	0	A	
Skill Writing	--	--	--	2	2	--	4	1	1
Pictorial Materials	--	--	--	1	1	1	1	--	--
Visual Displays	--	--	--	--	--	--	1	--	--
Verbal Communication	--	--	2	3	3	--	--	4	1
Non-precision Tools	--	--	--	1	1	--	--	--	--
Adjusting Equipment	--	--	--	1	1	--	1	--	--
Influencing Others	--	--	2	2	2	--	--	3	--
Exchanging Info	--	--	--	1	1	--	--	2	1
Good Precision	--	--	1	--	1	--	--	2	3
Attention - Details	--	--	--	2	2	2	3	--	--
Work Distractions	--	--	2	2	2	2	3	2	3
Keep up to Date	--	--	2	2	2	2	3	2	3
Total	0	0	7	10	17	6	3	14	1
									9

Table 6  
Zero-Order Validities of Experimental and Operational  
Variables for Global Performance

Predictor	Validity			
	EM (N=27) <sup>a</sup>	PN (N=54)	ST (N=37) <sup>a</sup>	UA (N=111)
<u>A Battery</u>				
Object Number	-.26	.13	-.03	-.01
Memory for Numbers (A)	-.15	.17	.22	.08
Counting Numbers	.03	.04	.42*	.06
Gestalt Completion	-.28	-.26*	.28	.06
Concealed Words	-.37*	-.14	.13	-.10
Hidden Patterns	-.04	.23	.33*	.11
Drift Direction	-.29	.07	.02	.06
Nonsense Syllogisms	-.30	.01	.30	-.06
Inference	.18	.19	.00	.13
<u>GRIP Battery</u>				
Memory for Objects	-.16	-.03	-.05	-.07
Memory for Words	-.33	.20	.13	.01
Memory for Numbers (V)	-.15	.20	.38*	-.01
Comparing Figures, Machine-paced	.02	-.10	.07	-.06
Comparing Figures, Self-paced	.06	.07	.21	.08
Recognizing Objects	-.11	-.06	.25	-.05
Memory for Patterns, True-False	.15	-.07	.42*	.07
Memory for Patterns, Free Response	.19	.21	.23	.19*
Twelve Questions	-.20	.28*	.21	.11
Password	.08	.13	.33*	.04
WORD-LAT	-.24	-.06	-.05	-.11
CLO-LAT	.05	.02	-.24	-.11
FIG-LAT	-.04	.00	.02	.04
PAT-ERR	-.24	-.17	-.26	-.13
<u>O Battery</u>				
AFQT	-.09	.15	-.12	.13
GCT	.01	.24	.11	.07
ARI	-.20	.10	.38*	.25**
MECH	.04	.23	-.04	.12
CLER	.21	-.15	.11	.19*
SONR	-.08	.15	-.08	-.03
RADO	-.06	.11	.15	.15
ETST	.16	.31	-.09	.33**
SHOP	.20	.38*	-.21	.17

<sup>a</sup>Complete data were not available for some of the tests.

\*Significant at  $P < .05$ .

\*\*Significant at  $P < .01$ .

Table 7

Multiple R's for Operational Battery only, and Operational  
Battery plus Experimental Battery, by Occupational Group

Group	Operational Battery only		Operational Battery plus O and A Variables	
	R	Shrunken Validity	R	Shrunken Validity
Electrician's Mate	.21	.00	.33	.00
Personnelman	.38	.12	.38	.12
Sonar Technician	.38	.22	.73	.58
Undifferentiated Apprenticeship	.33	.28	.41	.32

(STs), with the *A* battery being next most predictive. For recruits in the upper half (UA) of the recruit distribution, the *O* battery was the most predictive, with the *GRIP* battery being next.

Of the two tests in the *A* battery which had statistically significant validity coefficients for overall performance, one was a measure of Closure and the other measured Perceptual Speed. In contrast, three of the five *GRIP* measures which were significantly valid measured Short Term Memory and the other two were measures of Dealing with Concepts/Information. Significantly, neither the computerized measure of Closure (*Recognizing Objects*) nor its associated latency measure (CLO-LAT) was statistically significant for overall performance. The tests in the *O* battery which had statistically significant validity coefficients measured arithmetic reasoning, Perceptual Speed and electronic and mechanical knowledge.

The relatively high incidence of significant validities of tests in the *GRIP* and the *A* batteries for Sonar Technician appears to reflect the visual and perceptual requirements which have come to be associated with the ST job in recent years. Whereas formerly the machine representation of sonar echoes to Technicians was as auditory stimuli, on most of the equipment presently in use in the Navy the echoes are pres-

ented visually on cathode ray tube (CRT) displays. Apparently, it is the ability required for this important part of the Sonar Technician job which some *GRIP* tests were effective in measuring. The visual characteristics of the ST job also appear to explain the significant validity coefficient of *Counting Numbers*, the measure of Perceptual Speed in the *A* battery, in contrast to the non-significant coefficients of its counterparts in the *O* and the *GRIP* batteries (CLER and *Comparing Figures*). Scanning and searching tasks are not as prominent a part of the two latter tests as they are of *Counting Numbers*.

*Multiple regression.* Multiple regression statistics for optimal sets of the *O* and the *O* plus the experimental variables are shown in Table 7. Shrunken validity coefficients of the *O* battery were quite low—barely significant for most jobs—and the coefficients were lower for technical than for apprenticeship jobs. In contrast, the shrunken coefficients computed when the *A* and the *GRIP* variables were added to the predictor battery were quite high for the ST job, but failed to improve the predictability of EM and PN performance. The experimental variables provided an increase to the shrunken validity coefficient for STs of .36, compared with an increase of .04 for the Apprenticeship group.

### Discussion and Conclusions

The findings suggest that the frequently observed low positive correlations between scores on traditional paper-and-pencil tests and on-job performance result from the predictiveness of general mental ability for fundamental or basic skills required for on-job performance. Probably these skills consist primarily of general reasoning and learning abilities. On the other hand, for the specialized and demanding tasks associated with some technical jobs, tests measuring more specific and narrow abilities appear to be of greater value for predicting future job performance.

CRT equipment, in conjunction with a supporting computer, can be used to present and score on-line a number of tests which presently require hand scoring. Therefore, computerized equipment should make it feasible to include a greater variety of tests in personnel selection and classification batteries. Specifically, computerized testing equipment should make it feasible to include in batteries used for large-scale testing programs many tests which are presently excluded because of the cost and time requirements associated with hand scoring. Tests of this type in the present battery measured Short Term Memory and Closure.

In addition, the findings suggest that computerized tests were better predictors of the Dealing with Concepts/Information attribute than were their paper-and-pencil analogues. The fact that the computerized measures of this attribute significantly predicted on-job performance, whereas their paper-and-pencil counterparts did not, suggests that the computerized tests were more accurate for measuring the attribute as it was manifested in on-job performance.

Thus, computerized tests were more accurate than paper-and-pencil tests for measuring one of the five attributes covered in the present research. Extrapolation of this 20% rate of superiority of accuracy of computerized tests would imply that computerized tests would be found to be more accurate than paper-and-pencil tests for about 12 of the 63 variables identi-

fied by Mecham and McCormick (1969) but not included in the present research.

For the five abilities studied, the measurement characteristics of the *GRIP* tests as compared to their paper-and-pencil counterparts in the *A* battery, were as follows:

1. Computerized measurement and scoring of short-term recall using a variety of stimuli is feasible, and permits data collection and scoring to be carried out more easily and rapidly than is possible on paper-and-pencil tests measuring the same attribute. The findings of the present research suggest that tests of Short Term Memory are potentially valid measures of a job-relevant ability.
 

The number and type of significant validities for job elements and for overall performance varied from memory test to memory test, apparently to some extent related to the content of the test. *Memory for Numbers* and *Memory for Patterns* were the tests with the largest numbers of significant validity coefficients. Particularly useful as a predictor for job performance was *Memory for Patterns*, a test requiring exact memory of two-dimensional locations.
2. Measures of Perceptual Speed in the *A*, *O*, and *GRIP* batteries had high intercorrelations, and the job elements for which the validity coefficients of the *GRIP* tests were statistically significant were about the same as those for which the coefficients of its counterparts in the *A* and *O* batteries were statistically significant. However, the differences favored the paper-and-pencil measures, and the only significant validity coefficients with overall performance for tests of Perceptual Speed were for paper-and-pencil tests.

*Comparing Figures* in the *GRIP* battery employed only one of a number of possible approaches for utilizing the diverse capabilities of computerized equipment to measure Perceptual Speed. It is possible that other approaches with computerized tests may be

more successful for this purpose. However, the initial judgment concerning the usefulness of computerized tests for supplementing or replacing paper-and-pencil measurement of Perceptual Speed would be negative.

3. As is indicated by the validity coefficients of Closure tests in both the *A* and the *GRIP* batteries for both job elements and global performance, Closure tests measure an ability which is important for performance of Sonar Technicians. Interestingly, CLO-LAT, a straight latency measure (from the *GRIP* battery), substantially duplicated the validity pattern of *Hidden Patterns* (from the *A* battery) for job elements, although this similarity did not hold for the validity coefficients for overall performance.

In terms of their payoff as predictors of overall performance, computerized tests of Closure were not as useful as paper-and-pencil tests of Closure. However, the similarity of patterns of the computerized and paper-and-pencil tests for job elements as well as the apparent similarity of the format of these tests suggests that computerized tests for this attribute can be made to be as useful as paper-and-pencil tests.

4. The two experimental tests designed to measure Movement Detection were not closely related to one another and therefore did not provide evidence of a Movement Detection factor.
5. *GRIP* tests designed to measure the attribute of Dealing with Concepts/Information (*Twelve Questions* and *Password*) frequently correlated significantly with supervisory ratings of job element performance. Although the job elements for which *Twelve Questions* and/or *Password* had statistically significant validity coefficients were similar to those which were significant for their counterpart tests in the *O* battery (*Inference* and *Syllogistic Reasoning*), the *GRIP* measures had generally higher coefficients. Furthermore, only the

computerized tests had correlations with overall performance which were statistically significant. Thus, for measuring the Dealing with Concepts/Information attribute, computerized tests appear to offer significant improvements of accuracy over their paper-and-pencil analogues.

From the findings of the limited study described in the present article, it is clear that CRT equipment offers capabilities for increasing the predictiveness of personnel selection tests for on-job performance criteria. Furthermore, computerized tests appear to be especially good for measuring some job-relevant abilities. However, substantial additional research will be required to explicate the relative usefulness of computerized and paper-and-pencil tests for measurement not only of the 5 abilities covered in the present research but of 63 additional attributes identified by Mecham and McCormick (1969).

## References

- Cory, C. H., Rimland, B., & Bryson, R. A. Using computerized tests to measure new dimensions of abilities: An exploratory study. *Applied Psychological Measurement*, 1977, 1, 101-110.
- Darlington, R. B. Multiple regression in psychological research and practice. *Psychological Bulletin*, 1968, 69, 161-182.
- French, J. W., Ekstrom, R. B., & Price, L. A. *Manual for kit of reference tests for cognitive factors* (Rev. 1963). Princeton, NJ: Educational Testing Service, 1963.
- Lord, F. M. Some test theory for tailored testing. In W. H. Holtzman (Ed.), *Computer-assisted instruction, testing and guidance*. New York: Harper and Row, 1970.
- McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*, 1972, 57, 347-368.
- Mecham, R. C., & McCormick, E. J. *The rated attribute requirements of job elements in the Position Analysis Questionnaire*. (Report No. 1, AD 682490). Purdue University, Occupational Research Center, January 1969.

- Weiss, D. J. *Strategies of adaptive ability measurement*. (Research Report 74-5). Minneapolis: University of Minnesota, Department of Psychology Psychometric Methods Program, 1974. (AD A004270).
- Weiss, D. J., & Betz, N. E. *Ability measurement: Conventional or adaptive?* (Research Report 73-1). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, 1973. (AD 757788).
- Wood, R. Response-contingent testing. *Review of Educational Research*, 1974, 48, 529-544.
- Research under Project Order Number RR042-04 NR-150-335 and by the Navy Personnel Research and Development Center. Dr. Marshall Farr served as contract monitor for ONR. The opinions expressed are those of the author and do not necessarily reflect those of the Department of the Navy.*

#### **Author's Address**

Charles H. Cory, Navy Personnel Research and Development Center, San Diego, CA 92152.

#### **Acknowledgements**

*This research was supported by the Personnel and Training Research Programs of the Office of Naval*